

Page 1, line 22, please replace the paragraph with the following:

a3
[6] N. Shivakumar and H. Garcia-Molina, "SCAM: A Copy Detection Mechanism for Digital Documents," Proceedings of the Second International Conference in Theory and Practice of Digital Libraries, June 1995.

Page 1, line 25, please replace the paragraph with the following:

a4
[7] N. Shivakumar and H. Garcia-Molina, "Building a Scalable and Accurate Copy Detection Mechanism," Proceedings of Third International Conference on Theory and Practice of Digital Libraries, March 1996.

Page 2, line 4, please replace the paragraph with the following:

a5
[10] V. Chalana, A. Bruce, and T. Nguyen, "Duplicate Document Detection in DocBrowse," www.statsci.com/docbrowse/paper/spie98/node1.htm, July 31, 1999.

Page 2, line 6, please replace the paragraph with the following:

a6
[11] G. Salton, A. Wong, and C.S. Yang, "A Vector Space Model for Automatic Indexing," Comm. Of the ACM, vol. 18, no.11, pp. 613-620, November 1975.

Page 2, line 8, please replace the paragraph with the following:

a7
[12] M. F. Porter, "An Algorithm for Suffix Stripping," Program, vol. 14, no. 3, pp. 130-137, July 1980.

Page 2, line 13, please replace the paragraph with the following:

A8
Sub
C1
7

[14] R. S. Scotti and C. Lilly, "Analysis and Design of Test Corpora for Zero-Tolerance Government Document Review Process," Symposium for Document Image Understanding Technology, Annapolis, Maryland, April 1999, also reported at George Washington University Declassification Productivity Research Center, <http://dprec.seas.gwu.edu>, July 31, 1999.

Page 2, line 15, please replace the paragraph with the following:

A9

[15] D. Grossman, D. Holmes, and O. Frieder, "A Parallel DBMS Approach to IR in TREC-3", Overview of the Third Text Retrieval Conference (TREC-3), November 1994.